Enclosure 3

REPORT DOCUMENTATION PAGE	Form Approved OMB NO. 0704-0188

Public Reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comment regarding this burden estimates or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188,) Washington, DC 20503.

1. AGENCY USE ONLY (Leave Blank)	2. REPORT DATE July 2005	3. REPORT TYPE AND DATES COVERED **Manuscripts** **01 June 2004 – 31 December 2006**

4. TITLE AND SUBTITLE Opportunistic Link Scheduling, Power Control, and Routing for Multi-hop Wireless Networks over Time Varying Channels	5. FUNDING NUMBERS **W911NF0410224**
6. AUTHOR(S) Yih-Hao Lin and Rene L. Cruz	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **University of California. San Diego** **9500 Gilman Drive, La Jolla, CA 92093-0407**	8. PERFORMING ORGANIZATION REPORT NUMBER **N/A**
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) U. S. Army Research Office P.O. Box 12211 Research Triangle Park, NC 27709-2211	10. SPONSORING / MONITORING AGENCY REPORT NUMBER **N/A**

11. SUPPLEMENTARY NOTES

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision, unless so designated by other documentation.

12 a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release; distribution unlimited.	12 b. DISTRIBUTION CODE **N/A**

13. ABSTRACT (Maximum 200 words)

We consider a cross-layer optimization problem for multi-hop wireless networks over time varying channels. The system consists of L interfering links, where the transmission power and rate of link l (= 1, . . . , L) are specified in vectors $\vec{P} = [P_1 \ldots P_L]$ and $\vec{X} = [X_1 \ldots X_L]$ respectively. In every time slot, the scheduler schedules the transmissions by assigning a resource allocation vector $\vec{V} = [\vec{P}\ \vec{X}]^{\top}$. We denote the expectation of $\vec{V}$ by $E(\vec{V})$. Our objective is to find the optimal scheduling policy which minimizes the cost of average resource consumption while maintaining average service guarantees to each user. We develop a unified framework, in which the cost of the average resource consumption is given by a convex function $\vec{f}(E(\vec{V}))$ and the minimum average service guarantees are given by a set of convex constraints $\vec{g}(E(\vec{V})) \leq 0$. By means of convex optimization and stochastic approximation, we obtain the solution by solving the corresponding dual problem. An iterative algorithm is proposed and analyzed, which schedules the transmission powers and rates adapting to the channel variations. If the channel states are described by a finite-state mixing process, it is shown that our algorithm asymptotically attains the optimal cost.

14. SUBJECT TERMS scheduling, wireless networks, time varying channels, rate guarantees, duality, stochastic optimization	15. NUMBER OF PAGES 13
	16. PRICE CODE **N/A**

17. SECURITY CLASSIFICATION OR REPORT	18. SECURITY CLASSIFICATION ON THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	**UNCLASSIFIED**	**UNCLASSIFIED**	**U**

Opportunistic Link Scheduling, Power Control, and Routing for Multi-hop Wireless Networks over Time Varying Channels

Yih-Hao Lin and R. L. Cruz

Department of Electrical & Computer Engineering
University of California, San Diego
La Jolla, CA 92093-0407
y5lin@ucsd.edu, cruz@ece.ucsd.edu

Abstract

We consider a cross-layer optimization problem for multi-hop wireless networks over time varying channels. The system consists of L interfering links, where the transmission power and rate of link l $(= 1, \ldots, L)$ are specified in vectors $\vec{P} = [P_1 \ldots P_L]$ and $\vec{X} = [X_1 \ldots X_L]$ respectively. In every time slot, the scheduler schedules the transmissions by assigning a resource allocation vector $\vec{V} = [\vec{P}\ \vec{X}]^{\mathsf{T}}$. We denote the expectation of $\vec{V}$ by $E(\vec{V})$. Our objective is to find the optimal scheduling policy which minimizes the cost of average resource consumption while maintaining average service guarantees to each user. We develop a unified framework, in which the cost of the average resource consumption is given by a convex function $f(E(\vec{V}))$ and the minimum average service guarantees are given by a set of convex constraints $\vec{g}(E(\vec{V})) \leq 0$. By means of convex optimization and stochastic approximation, we obtain the solution by solving the corresponding dual problem. An iterative algorithm is proposed and analyzed, which schedules the transmission powers and rates adapting to the channel variations. If the channel states are described by a finite-state mixing process, it is shown that our algorithm asymptotically attains the optimal cost.

1 Introduction

The growing interest in multi-hop wireless network increases the demand for efficient resource allocation methods. To maintain the quality of service and manage the cost of average resource consumption, a scheduler is responsible for coordinating all communicating nodes in the system. In this paper we assume that the scheduler is centralized. This allows us to obtain insight into the structure of the problem, and our solution to the problem in this case serves as a benchmark for distributed scheduling algorithms.

In each time slot, a scheduler specifies the transmission power and rate on each link. Since wireless channel conditions are time varying, scheduling a transmission over a bad channel results in high power consumption and may result in large interference. Therefore, an intelligent scheduler should take into account the instantaneous *channel state information* (CSI), such as background noise, pathloss gain, and fading, as well the variations of these quantities in time and space, in order to achieve an efficient resource utilization.

Several studies have addressed the problem of throughput utility maximization, considering time varying channels. In [9], Tse proposed the *proportional fair scheduler* (PFS), which exploits multi-user diversity in a time varying environment. Under the stationary assumptions on the channels, the throughput vector $[X_1^{avg}, \dots, X_L^{avg}]$ of PFS achieves the maximum of the logarithmic utility function $\sum_{l=1}^{L} \log(X_l^{avg})$ among all schedulers [5]. Agrawal [1] and Stolyer [8] further generalize this idea by considering arbitrary utility functions of the form $\sum_{l=1}^{L} f_l(X_l^{avg})$, where $f_l(\cdot)$ is concave and differentiable.

One weakness of the utility maximizing scheduling is that the optimal scheduler favors users with better channel conditions. This implicitly decreases the priority of users with poor channel quality. However, it may be highly desirable to impose a minimum average performance requirement on each link. To address this issue, Lee et al. [7] reconsidered the throughput maximization problem and impose a minimum throughput constraint on each user. However, their solution is only asymptotically optimal in the limit of an infinite number of users.

The structure of optimal scheduler under time varying channels is investigated in this paper, in which we minimize the cost of average resource consumption while maintaining a minimum average performance guarantee on each link. The problem is formulated as an optimization. An iterative algorithm is proposed to obtain the solution. By means of convex optimization and stochastic approximation, we verify that our algorithm approaches the optimum irrespective of the number of users. We show that besides the link layer scheduling and power control problem, our framework can be further extended to solve other network layer optimization problems, such as power efficient routing.

2 Notations and System Model

We consider an omni-antenna wireless system consisting of N nodes and L links. The instantaneous transmission power and rate are specified in the vectors $\vec{P} = [P_1 \dots P_L]$ and $\vec{X} = [X_1 \dots X_L]$ respectively. Time is divided into fixed duration intervals, called time slots. At the beginning of time slot k $(= 1, \dots, \infty)$, the scheduler specifies the resource allocation for the system in the vector $\vec{V} = [\vec{P}\ \vec{X}]^{\mathsf{T}}$. We model the underlying time varying phenomena as a stationary process $\{\omega(k), k \geq 0\}$ with finite state space S and stationary distribution $\{s(\omega), \omega \in \mathcal{S}\}$. The outcome of the process is called the system state. In state ω, the power loss gain from the transmitter of link l_1 to the receiver of link l_2 is denoted by $G_{l_1 l_2}(\omega)$ including path loss and fading. On link l, the receiving end experiences noise power $\xi_l(\omega)$. We hence have the *signal to interference and noise ratio* (SINR) $\gamma_l(\omega) = \frac{G_{l,l}(\omega) P_l}{\sum_{k \neq l} G_{k,l}(\omega) P_k + \xi_l(\omega)}$. The sets of all outgoing links and incoming links at node n are denoted by $\mathcal{E}(n)$ and $\mathcal{F}(n)$ respectively. Assuming the maximum instantaneous data rate possible by link l is a function $R(\gamma_l)$ of SINR and the peak transmission power of node n is limited to P_n^{max}, a resource allocation $\vec{V}$ is feasible if and only if $\vec{X}$ and $\vec{P}$ satisfy the following constraints.

(**Power Constraints**)

$$\begin{array}{ll} P_l \geq 0 & \text{for } l = 1, \dots, L \\ \sum_{l \in \mathcal{E}(n)} P_l \leq P_n^{max} & \text{for } n = 1, \dots, N \end{array} \tag{1}$$

(**Rate Constraints**)

$$\begin{array}{ll} X_l \geq 0 & \text{for } l = 1, \dots, L \\ X_l \leq R(\gamma_l(\omega)) & \text{for } l = 1, \dots, L \end{array} \tag{2}$$

We denote $\mathcal{D}(\omega)$ as the set of all feasible resource allocation $\vec{V}$. It is uniformly bounded over $\mathcal{S}$ under the finite state assumption.

3 Scheduling Policy and Achievable Set of Average Resource Allocation

In this paper, we consider a probabilistic scheduling policy, where the scheduling decisions are random variables over the state space $\mathcal{S}$. Specifically, in state ω, a scheduling policy π specifies a feasible resource allocation $\vec{V}$ randomly from a specified control set $\mathcal{K}(\omega)$ with probability distribution $\phi_\omega(\vec{V})$, where $\mathcal{K}(\omega)$ is a subset of $\mathcal{D}(\omega)$, $\phi_\omega(\vec{V}) \geq 0$ and $\sum_{\vec{V}\in\mathcal{K}(\omega)} \phi_\omega(\vec{V}) = 1$. Therefore, given a scheduling policy π, the average resource allocation $E_\pi(\vec{V})$ can be evaluated as below.

$$E_\pi(\vec{V}) \;\; = \;\; \sum_{\omega\in\mathcal{S}} s(\omega) \sum_{\vec{V}\in\mathcal{K}(\omega)} \phi_\omega(\vec{V})\, \vec{V}. \tag{3}$$

We omit the subscript π when the policy is known explicitly. We define the achievable set $\tilde{\mathcal{D}} = \left\{\vec{v} \middle| \text{ There exists } \pi \text{ such that } \vec{v} = E_\pi(\vec{V})\right\}$. The uniform boundedness of $\mathcal{D}(\omega)$ implies the boundedness of $\tilde{D}$. Moreover, one can verify the convexity of $\tilde{\mathcal{D}}$ by a time-sharing argument. In the next section, we introduce the framework for optimal link scheduling and power control, where we only look into link layer scheduling problems. The extensions to network layer routing problems are discussed in section 5.

4 Generalized Framework for Link Scheduling and Power Control

Considering link scheduling and power control, most of the link layer resource management problems can be formulated as a special case of the following optimization,

$$\begin{aligned} \text{minimize} \quad & f(E(\vec{V})) \\ \text{subject to} \quad & \vec{g}(E(\vec{V})) \leq 0 \\ & E(\vec{V}) \in \tilde{\mathcal{D}}, \end{aligned} \tag{4}$$

where $f(\cdot)$ is a scalar valued convex function and $\vec{g}(\cdot)$ is a vector valued convex function. For example, if we set the cost function to $f(E(\vec{V})) = -\sum_{l=1}^{L} \log(E(X_l))$, the optimization reduces to the proportional fair scheduling problem. If we let $f(E(\vec{V})) = \sum_{l=1}^{L} E(P_l)$, it becomes the power efficient scheduling problem.

Since the constraint set $\tilde{\mathcal{D}}$ is bounded and convex, if $E(\vec{V})$ is considered as the independent variable, the problem (4) fits in with the framework of convex optimization. The duality technique [2] in convex analysis helps to simplify the problem by moving the constraints into the cost function. The dual function $q(\vec{\beta})$ is defined as

$$q(\vec{\beta}) \;\; = \;\; \min_{E(\vec{V})\in\tilde{\mathcal{D}}} L(E(\vec{V}), \vec{\beta}), \tag{5}$$

where $L(E(\vec{V}), \vec{\beta}) := f(E(\vec{V})) + \vec{\beta}\cdot\vec{g}(E(\vec{V}))$ is the Lagrangian of the primal problem. The dot product is defined as the inner product of two vectors. One can verify that $q(\vec{\beta})$

is a concave function [2]. The duality theory defines the dual problem below.

$$\begin{aligned} \text{maximize} \quad & q(\vec{\beta}) \\ \text{subject to} \quad & \vec{\beta} \geq 0 \end{aligned} \tag{6}$$

The solution of (6) is denoted by $\vec{\beta}^*$, where the elements of the vector $\vec{\beta}^*$ are called Lagrange multipliers. From the weak duality theorem, it is recognized that the dual optimal value q^* is a lower bound of the primal optimal value f^*. If there exists a feasible solution in the relative interior of $\tilde{\mathcal{D}}$, the strong duality theorem ensures that $q^* = f^*$. In the following subsections we first solve the problem with a linear cost function and linear constraints, and then generalize it to nonlinear problems with the help of additional auxiliary variables.

4.1 Linear Cost and Linear Constraints

For a system with a linear cost function $f(\vec{v}) = \vec{\alpha} \cdot \vec{v}$ and constraints $\vec{g}(\vec{v}) = \mathbf{A}\vec{v} - \vec{b}$, the primal problem (4) becomes

$$\begin{aligned} \text{minimize} \quad & \vec{\alpha} \cdot E(\vec{V}) \\ \text{subject to} \quad & \mathbf{A}E(\vec{V}) \leq \vec{b} \\ & E(\vec{V}) \in \tilde{\mathcal{D}}, \end{aligned} \tag{7}$$

where $\vec{\alpha}$ and $\vec{b}$ are vectors and $\mathbf{A}$ is a matrix.

Theorem 1. *Given the optimal dual solution $\vec{\beta}^*$, the optimal scheduling policy π^* for (7) is given by $\{ \ (\mathcal{K}^*(\omega), \phi^*_\omega(\vec{V}) \ ) \mid \omega \in \mathcal{S} \ \}$, the optimal control sets and the corresponding distributions, where $\mathcal{K}^*(\omega) = \arg\min_{\vec{V} \in \mathcal{D}(\omega)} \left[\vec{\alpha} \cdot \vec{V} + \vec{\beta}^* \cdot (\boldsymbol{A}\vec{V} - \vec{b}) \right]$, and $\phi^*_\omega(\vec{V})$ is the distribution over $\mathcal{K}^*(\omega)$ such that $E_\pi(\vec{V}) \in \tilde{\mathcal{D}}$, $\vec{g}(E_\pi(\vec{V})) \leq 0$, and $\vec{\beta}^* \cdot \vec{g}(E_\pi(\vec{V})) = 0$.*

Proof. The dual function of (7) can be written as

$$\begin{aligned} q(\vec{\beta}) &= \min_{E(\vec{V}) \in \tilde{\mathcal{D}}} \left\{ \vec{\alpha} \cdot E(\vec{V}) + \vec{\beta} \cdot (\mathbf{A}E(\vec{V}) - \vec{b}) \right\} \qquad (8) \\ &= \min_{\pi} \sum_{\omega \in \mathcal{S}} s(\omega) \sum_{\vec{V} \in \mathcal{K}(\omega)} \phi_\omega(\vec{V}) \left[\vec{\alpha} \cdot \vec{V} + \vec{\beta} \cdot (\mathbf{A}\vec{V} - \vec{b}) \right] \\ &= \sum_{\omega \in \mathcal{S}} s(\omega) \min_{\mathcal{K}(\omega), \phi_\omega(\vec{V})} \Big\{ \sum_{\vec{V} \in \mathcal{K}(\omega)} \phi_\omega(\vec{V}) \left[\vec{\alpha} \cdot \vec{V} + \vec{\beta} \cdot (\mathbf{A}\vec{V} - \vec{b}) \right] \Big\} \qquad (9) \end{aligned}$$

To achieve the minimum in (9), the control set $\mathcal{K}(\omega)$ must be a subset of

$$\arg \min_{\vec{V} \in \mathcal{D}(\omega)} \left[\vec{\alpha} \cdot \vec{V} + \vec{\beta} \cdot (\mathbf{A}\vec{V} - \vec{b}) \right]. \tag{10}$$

The the proof is accomplished by the following lemma. □

Lemma 1. *(Optimality Conditions for Convex Optimization) [2]. The primal and dual variable pair, $(E(\vec{V}^*), \vec{\beta}^*)$, is an optimal solution-Lagrange multiplier pair of (4) if and only if*

$E(\vec{V}^*) \in \tilde{\mathcal{D}}, \ \vec{g}(E(\vec{V}^*)) \leq 0$	*(Primal Feasibility)*
$\vec{\beta}^* \geq 0,$	*(Dual Feasibility)*
$E(\vec{V}^*) = \arg\min_{E(\vec{V}) \in \tilde{\mathcal{D}}} L(E(\vec{V}), \vec{\beta}^*)$	*(Lagrangian Optimality)*
$\vec{\beta}^* \cdot \vec{g}(E(\vec{V}^*)) = 0$	*(Complementary Slackness)*

Since $\{\phi_\omega(\vec{V})\}$ sum up to one, from (9), we can further deduce that

$$\begin{aligned} q(\vec{\beta}) &= \sum_{\omega\in\mathcal{S}} s(\omega) \min_{\vec{V}\in\mathcal{D}(\omega)} \left[\vec{\alpha}\cdot\vec{V} + \vec{\beta}\cdot(\mathbf{A}\vec{V}-\vec{b})\right] \\ &= E\left[\min_{\vec{V}\in\mathcal{D}(\omega)} \left\{\vec{\alpha}\cdot\vec{V} + \vec{\beta}\cdot(\mathbf{A}\vec{V}-b)\right\}\right]. \end{aligned} \tag{11}$$

For ease of exposition, we define the quasi-dual function as

$$q(\omega,\vec{\beta}) := \min_{\vec{V}\in\mathcal{D}(\omega)} \left\{\vec{\alpha}\cdot\vec{V} + \vec{\beta}\cdot(\mathbf{A}\vec{V}-b)\right\} \tag{12}$$

$$= \vec{\alpha}\cdot\vec{V}^*(\omega,\vec{\beta}) + \vec{\beta}\cdot(A\vec{V}^*(\omega,\vec{\beta})-\vec{b}), \tag{13}$$

where $\vec{V}^*(\omega,\vec{\beta})$ is a solution to the minimization on the right hand side of (12). Simple convexity arguments ensure that $q(\omega,\vec{\beta})$ is concave in $\vec{\beta}$. Substituting $q(\omega,\vec{\beta})$ into (11) and (6), the dual problem is turned into the following stochastic convex optimization problem.

$$\begin{aligned} \text{maximize} \quad & E[q(\omega,\vec{\beta})] \\ \text{subject to} \quad & \vec{\beta}\geq 0 \end{aligned} \tag{14}$$

For any concave function $q(\vec{\beta})$, a vector $\vec{\mu}$ is called a subgradient of $q(\vec{\beta})$ if it satisfies the inequality $q(\vec{\beta}')-q(\vec{\beta})\leq\vec{\mu}\cdot(\vec{\beta}'-\vec{\beta})$ for all $\vec{\beta}'$. By Danskin's theorem [2], one can show that the vector $(\mathbf{A}\vec{V}^*(\omega,\vec{\beta})-\vec{b})$ is a subgradient of $q(\omega,\vec{\beta})$ with respect to $\vec{\beta}$.

To solve (14), we apply the following algorithm.

Algorithm 1. *(Quasi-Gradient Method [3]) At the beginning of the k^{th} time slot, update the dual variables $\vec{\beta}^k$ by the recursion (15) below with step size ϵ_k, where $\epsilon_k>0$, $\sum_k\epsilon_k=\infty$, and $\sum_k\epsilon_k^2<\infty$.*

$$\vec{\beta}^k = [\vec{\beta}^{k-1}+\epsilon_{k-1}\vec{\mu}^{k-1}]^+, \qquad (\vec{\beta}^0=0,\ \vec{\mu}^0=0) \tag{15}$$

The vector $\vec{\mu}^k$ is a subgradient of $q(\omega(k),\vec{\beta}^k)$ and $[\ \cdot\ ]^+ := \max(\cdot,0)$. We choose the vector $\vec{\mu}^k=(\boldsymbol{A}\vec{V}^*(k)-\vec{b})$ in particular, where

$$\vec{V}^*(k)\in\mathcal{K}^*(\omega(k)) = \arg\min_{\vec{V}\in\mathcal{D}(\omega(k))} \left\{\vec{\alpha}\cdot\vec{V}+\vec{\beta}^k\cdot(\boldsymbol{A}\vec{V}-b)\right\}.$$

If $\{\omega(k),k\geq 0\}$ are *independent identically distributed* (i.i.d.), the algorithm ensures that $\vec{\beta}^k\to\vec{\beta}^*$ as $k\to\infty$ with probability one. In addition, if we choose the step size $\epsilon_k=a_k/k$, where $\lim_k a_k\to a$ and $a>0$, the long-term average of $\vec{V}^*(k)$ converges to the primal optimal solution of (7) with probability one.

Remark 1. *The i.i.d. assumption on $\omega(k)$ can be weaker. Algorithm 1 is still applicable if the dependency among $\{\omega(k),k\geq 0\}$ decreases in time. When the dual solution $\vec{\beta}^*$ is unique, one can verify via the stochastic approximation ([4], [6] Theorem 8.2.5 and 5.2.2) that $\vec{\beta}^k\to\vec{\beta}^*$ with probability one if $\omega(k)$ satisfies the mixing property: $|E_k\mu(\vec{\beta},\omega(i))-E\mu(\vec{\beta},\omega(i))|\to 0$ as $i\to\infty$, where E_k denotes the expectation conditioned on $\mathcal{F}_k$, the filtration of $\{\vec{\beta}^1,\omega(1),\ldots,\vec{\beta}^k,\omega(k)\}$.*

To prove the asymptotic optimality, we rewrite (15) as

$$\vec{\beta}^k = \vec{\beta}^{k-1} + \frac{a_k}{k}(\vec{\mu}^{k-1} + \vec{z}^{k-1}) \ , \tag{16}$$

where $\vec{z}^k$ is the vector of reflection terms of the $[\,\cdot\,]^+$ operator and $\vec{z}^k \geq 0$. Considering a telescoping sum on both sides, we derive the following.

$$\begin{aligned} \vec{\beta}^t &= \vec{\beta}^{t-1} + \frac{1}{t}(a_t\vec{\mu}^{t-1} + a_t\vec{z}^{t-1} + \vec{\beta}^{t-1} - \vec{\beta}^{t-1}) \qquad (17) \\ &= \frac{1}{t}\sum_{k=0}^{t-1}(a_k - a)\vec{\mu}^k + \frac{1}{t}\sum_{k=0}^{t-1}(a_k - a)\vec{z}^k + \frac{1}{t}\sum_{k=0}^{t-1} a\vec{\mu}^k + \frac{1}{t}\sum_{k=0}^{t-1} a\vec{z}^k + \frac{1}{t}\sum_{k=0}^{t-1}\vec{\beta}^k \end{aligned}$$

Since $\vec{\beta}^k \to \vec{\beta}^*$ with probability one, we have $\lim_{k\to\infty}\frac{1}{t}\sum_{k=0}^{t-1}\vec{\beta}^k = \vec{\beta}^*$ with probability one. The boundedness of $\mathcal{D}(\omega)$ implies that $\vec{\mu}^k = (\mathbf{A}\vec{V}^*(k) - \vec{b})$ is bounded. Because $\vec{z}^k$ is the reflection term, it is also bounded. Note that $\lim_{k\to\infty}(a_k - a) = 0$. Taking limits on (17) and then subtracting $\vec{\beta}^*$ from both sides, we arrive at

$$\lim_{k\to\infty}\left(\frac{1}{t}\sum_{k=0}^{t-1}\vec{\mu}^k + \frac{1}{t}\sum_{k=0}^{t-1}\vec{z}^k\right) = 0 \quad \text{with prob. 1} \tag{18}$$

When β_m^*, the m^{th} entry of $\vec{\beta}^*$, is positive, the corresponding reflection term z_m^k vanishes for k large enough. By (18), we therefore have $\lim_{k\to\infty}\frac{1}{k}\sum_{t=0}^{k-1}\mu_m^t = 0$. Similarly, if $\beta_m^* = 0$, one can show that $\limsup_{k\to\infty}\frac{1}{k}\sum_{t=0}^{k-1}\mu_m^t \leq 0$. Replacing $\vec{\mu}^k$ with $(\mathbf{A}\vec{V}^*(k)-\vec{b})$, we get the asymptotic feasibility described in the theorem below.

Theorem 2. *(Feasibility) If the limiting time average of $\vec{V}^*(k)$ exists, it lies in the feasible region of (7) with probability one, which means*

$$\boldsymbol{A}\left[\lim_{t\to\infty}\frac{1}{t}\sum_{k=0}^{t-1}\vec{V}^*(k)\right] - \vec{b} \leq 0 \quad \textit{with prob. 1.}$$

To show the asymptotic optimality, we take time average on (13) and then group the summands with respect to system state ω. The theorem below is then derived by applying the stationary property of $\omega(k)$.

Theorem 3. *(Optimality) If the limiting time average of $\vec{V}^*(k)$ exists, it attains the optimal solution with probability one. That is*

$$f^* = \vec{\alpha}\cdot\left\{\lim_{t\to\infty}\frac{1}{t}\sum_{k=0}^{t-1}\vec{V}^*(k)\right\} \quad \textit{with prob. 1.}$$

4.2 Convex Cost or Convex Constraints

If $f(\cdot)$ or $\vec{g}(\cdot)$ are nonlinear functions, we cannot apply the decomposition in (9) directly using the separable property. To accommodate the non-separable issue, we define an

auxiliary variable $\vec{Y}$ to substitute $E(\vec{V})$ in f and $\vec{g}$. Assuming the elements of $\tilde{\mathcal{D}}$ are bounded above componentwise by a vector $\vec{D}^{max}$, the optimization (4) becomes

$$\begin{aligned} \text{minimize} \quad & f(\vec{Y}) \\ \text{subject to} \quad & g(\vec{Y}) \le 0, \quad 0 \le \vec{Y} \le \vec{D}^{max} \\ & \vec{Y} = E(\vec{V}), \quad E(\vec{V}) \in \tilde{\mathcal{D}} \end{aligned} \tag{19}$$

By weak duality theory and convexity, the solution can be approached using the algorithm similar to that for the linear cases. Note that if $\vec{D}^{max}$ is not known explicitly, one can use an estimate instead and update it in each time slot.

5 Optimal Routing and Flow Rate Assignments

Now we extend our framework to consider other network layer optimizations. There are J flows in the system indexed by $j = 1, \ldots, J$. The source and destination nodes of flow j are denoted by $n_s(j)$ and $n_d(j)$ respectively. All links used by flow j for routing form a set $\mathcal{G}(j)$ of size m_j. The collection of all flows on link l constitutes the set $\mathcal{H}(l)$. In a time slot, link l has data rate C_{lj} for flow $j \in \mathcal{H}(l)$. Given $\mathcal{G}(j) = \{l_1, \ldots, l_{m_j}\}$, we define $\vec{C}_j = [C_{l_1 j} \ \ldots \ C_{l_{m_j} j}]$ as the routing vector of flow j. Furthermore, we group all routing vectors into a vector, $\vec{C} = [\vec{C}_1 \ \ldots \ \vec{C}_J]$, called the flow vector. The essential requirement for routing problem is to maintain minimum end-to-end throughput $\tilde{C}_j$ on each flow j. Under this requirement, two constraints are imposed on $\{C_{lj}\}$.

(**Flow Constraints**)

$$\begin{aligned} & \textstyle\sum_{j \in \mathcal{H}(l)} C_{lj} \le X_l \quad , l = 1, \ldots, L \\ & C_{lj} \ge 0 \end{aligned} \tag{20}$$

(**Flow Conservation Constraints**)
For $n = 1, \ldots, N. \quad j = 1, \ldots, J$

$$\sum_{l \in \mathcal{E}(n),\, l \in \mathcal{G}(j)} E(C_{lj}) - \sum_{l \in \mathcal{F}(n),\, l \in \mathcal{G}(j)} E(C_{lj}) + \nu_{nj} = 0, \quad \text{where} \quad \nu_{nj} = \begin{cases} \tilde{C}_j, & \text{if } n = n_d(j) \\ -\tilde{C}_j, & \text{if } n = n_s(j) \\ 0, & \text{otherwise} \end{cases} \tag{21}$$

The resource allocation vector now becomes $\vec{V} = [\vec{P} \ \vec{X} \ \vec{C}]^{\mathsf{T}}$. In particular, the new feasible set $\mathcal{D}(\omega)$ consists of all vectors $\vec{V} = [\vec{P} \ \vec{X} \ \vec{C}]^{\mathsf{T}}$ satisfying the power constraints, rate constraints and flow constraints. In the following subsection, we discuss our framework for the power efficient routing problem.

5.1 Power Efficient Routing

The power efficient routing problems can be formulated as the optimization below.

$$\begin{aligned} \text{minimize} \quad & \sum_{l=1}^{L} E(P_l) \\ \text{subject to} \quad & E[\vec{V}] \ \text{ satisfies (21)} \\ & E(\vec{V}) \in \tilde{\mathcal{D}} \end{aligned} \tag{22}$$

The quasi-dual function of this problem is given below.

$$q(\omega, \vec{\beta}) = \min_{\vec{V} \in \mathcal{D}(\omega)} \sum_{l=1}^{L} P_l + \sum_{j=1}^{J} \sum_{n=1}^{N} \beta_{kj} \Big(\sum_{l \in \mathcal{E}(n),\, l \in \mathcal{G}(j)} C_{lj} - \sum_{l \in \mathcal{F}(n),\, l \in \mathcal{G}(j)} C_{lj} + \nu_{nj} \Big) \tag{23}$$

The dual problem is therefore expressed as

$$\begin{array}{ll} \text{maximize} & Eq(\omega, \vec{\beta}) \\ \text{subject to} & \vec{\beta} \in \mathbb{R}^m, \end{array} \tag{24}$$

where m is the dimension of the constraints. To evaluate (23), first, we rearrange the summands on the right hand side of (23) with respect to C_{lj},

$$q(\omega, \vec{\beta}) = \min_{\vec{V} \in \mathcal{D}(\omega)} \sum_{l=1}^{L} P_l + \sum_{j=1}^{J} \sum_{l=1}^{L} \sigma_{lj} C_{lj} \ + \ \sigma_0, \quad \text{where} \ \ \sigma_0 = \sum_{j=1}^{J} \sum_{n=1}^{N} \beta_{kj} \nu_{nj} \tag{25}$$

and σ_{lj} is the new coefficient of C_{lj} after rearrangement. For fixed $\vec{P}$ and $\vec{X}$, equation (25) is a linear optimization problem in C_{lj} and the flow constraints form a polytope for $\{C_{lj}\}$. Hence it has solutions at extreme points of (20). One can verify that the optimal flow rates borne on each link are given by

$$C_{lj}^* = \begin{cases} X_l & \text{if } \sigma_{lj} = \arg\min_{j \in \mathcal{H}(l)} \sigma_{lj}, \text{ and } \sigma_{lj} < 0 \\ & \text{(If not unique, pick up one randomly)} \\ 0 & \text{else.} \end{cases} \tag{26}$$

Substituting (26) into (25) and rearranging the summands according to X_l, we arrive at the equation

$$q(\omega, \vec{\beta}) = \min_{\vec{V} \in \mathcal{D}(\omega)} \sum_{l=1}^{L} P_l + \sum_{l=1}^{L} \sigma'_l X_l + \sigma_0, \tag{27}$$

where σ'_l is the corresponding coefficient after the second rearrangement. Similarly, fixing $\vec{P}$, the optimal transmission rate on link l is given by

$$X_l^* = \begin{cases} R_l(\gamma_l) & \text{if } \sigma'_l < 0 \\ 0 & \text{if } \sigma'_l \geq 0. \end{cases} \tag{28}$$

Substituting (28) into (27) and then grouping the summands according to $R_l(\gamma_l)$, the evaluation of quasi-dual function reduces to the following optimization with respect to transmission powers:

$$q(\omega, \vec{\beta}) = \min_{\vec{V} \in \mathcal{D}(\omega)} \sum_{l=1}^{L} P_l + \sum_{l=1}^{L} \sigma''_l R_l(\gamma_l) + \sigma_0,$$

where σ''_l is the corresponding coefficient. The solution of (22) then can be obtained through the following algorithm.

Algorithm 2. *At time slot k, the optimal scheduling policy schedules a transmission power vector $\vec{P}^*(k)$ from the set below.*

$$\arg \min_{\vec{V} \in \mathcal{D}(\omega(k))} \left\{ \sum_{l=1}^{L} P_l + \sum_{l=1}^{L} \sigma''_l R_l(\gamma_l) \right\} \tag{29}$$

The optimal data rate $X_l^*(k)$ and flow rate $C_{lj}^*(k)$ can be deduced from (28) and (26). The dual variable is updated recursively as

$$\vec{\beta}^{k+1} = \vec{\beta}^k + \frac{a_{k+1}}{k+1}\left[\sum_{l\in\mathcal{E}(n)} C_{lj}^*(k) - \sum_{l\in\mathcal{F}(n)} C_{lj}^*(k) + \nu_{nj}\right], \qquad (\ \vec{\beta}^0 = 0\ )$$

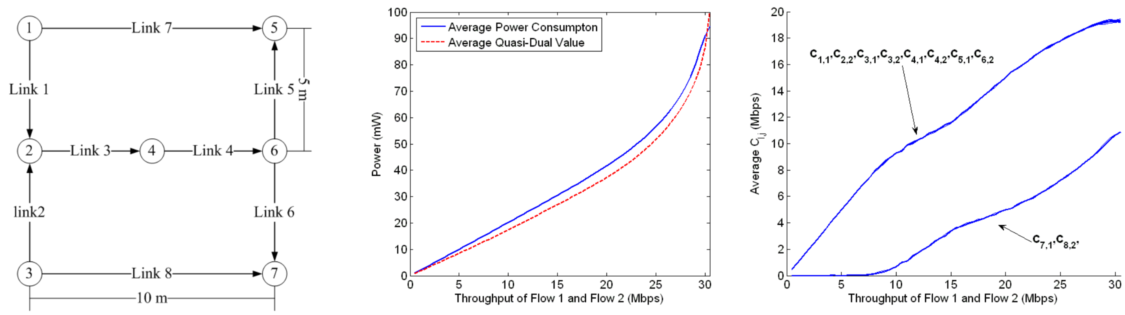
where $\lim_{k\to\infty} a_k = a$ *and* $a > 0$.

Let $\vec{V}^*(k) = [\vec{P}^*(k)\ \vec{X}^*(k)\ \vec{C}^*(k)]^\mathsf{T}$, if the process $\omega(k)$ is i.i.d. or mixing as described before, the long-term average of $\vec{V}^*(k)$ converges to the optimal solution of (22).

In the next section, we exam a numerical example of power efficient routing, where we apply the linear rate function $R(\gamma_l) = W'\gamma_l$. The cost function in the evaluation of (29) then becomes componentwise concave in P_l. One can show that the solutions to such problems happen at the extreme points of power constraints (1).

6 Numerical Examples

This example consists of 7 nodes and 8 links, which are depicted in figure 1(a) . There are two flows, each requires minimum throughput $\tilde{C}$. Flow 1 originates at node 1 and is destined to node 5; it exploits links 1, 3, 4, 5, and 7 to route the traffic. Flow 2 originates at node 3 and is destined to node 7; it uses link 2, 3, 4, 6, and 8 to route the traffic. We set P_n^{max} to 50 mW and W' to 50 MHz. The channel states are i.i.d. The background noises are normally distributed with mean zero and an average power of 1 mW. Moreover, the noise power ξ_n is truncated above at 10 mW and rounded to the second digit in precision. The channel gain $G_{l_1 l_2}$ is given by $e/d^2(l_1, l_2)$, where e is an exponential random variable with mean 1 and $d(l_1, l_2)$ is the distance between the transmitter of link l_1 and the receiver of link l_2. In addition, the factor e is truncated from above at 2 and rounded to the third digit in precision. The step size is $\epsilon_k = \frac{2.5}{500+k}$. To investigate how system performs with respect to $\tilde{C}$, we gradually increase the minimum throughput requirements from 0.5 to 30 Mbps. The traces of the average flow rates carried on each link are plotted in figure 1(c). The upper curve contains 8 overlapped traces of the time averages of $\{C_{11}, C_{22}, C_{31}, C_{32}, C_{41}, C_{42}, C_{51}, C_{62}\}$ and the lower curve 2 overlapped traces of the time averages of $\{C_{71}, C_{82}\}$. Because the topology is symmetric, we only focus on the behavior of flow 1. It is recognized that when the requested throughput is below 5 Mbps, the optimal route of flow 1 is $\{1 \to 2 \to 4 \to 6 \to 5\}$; as the requested throughput increases, the direct route $\{1 \to 5\}$ comes into play. Although both $\{1 \to 2 \to 4 \to 6 \to 5\}$ and $\{1 \to 5\}$ are energy efficient paths, on route $\{1 \to 2 \to 4 \to 6 \to 5\}$ each link contributes less interference to the system, and the possibility that all of the links are in deep fade at the same time is small. In figure 1(b), we gather the data from slot 5001 to slot 25000 and plot the average power consumption and average quasi-dual value respect to $\tilde{C}$. We can see that the curve of average power consumption closely follows that of the average $q(\omega, \beta_k)$ as expected. The gap between both curves will diminish to zero in the limit of an infinite number of time slots. When $\tilde{C}$ is small, the scheduler works like a TDMA system. In other words, in every slot, at most one of the links is activated, and the scheduler refrains from transmitting on a link unless that channel is in good condition. Therefore, the total power consumption increases linearly in the low throughput region. However, as $\tilde{C}$ goes up to 20Mbps, more links need to be activated in a slot. The total power consumption then increases nonlinearly due to interference.



(a) Network Topology (b) Power Consumption (c) Average Flow Data Rates

Figure 1:

7 Conclusions

In this paper, we have shown how a joint link scheduling, power control, and routing optimization problem over time varying channels can be formulated as a stochastic convex optimization. In addition, we proposed a centralized iterative algorithm which exploits the CSI to schedules the transmissions. We have proved that the optimality and feasibility is attained asymptotically. The numerical results show that for flows with minimum throughput requirements, to save power consumption the scheduler opportunistically routes the traffic through links with better channel conditions.

References

[1] R. Agrawal and V. Subramanian. "Optimality of Certain Channel Aware Scheduling Policies". In *Proc. 2002 Allerton Conf.*, pages 1533–1542. Univ. of Illinois Press.

[2] D.P. Bertsekas. *Nonlinear Programming.* Athena Scientific, 2nd edition, 2001.

[3] Yu. Ermoliev. *Numerical Techniques for Stochastic Optimization.* Springer-Verlag, 1988.

[4] Y. He, M.C. Fu, and S.I. Marcus. "Convergence of Simultaneous Perturbation Stochastic Approximation for Nondifferentiable Optimization". *IEEE Trans. Automat. Contr.*, 48(8):1459, Aug. 2003.

[5] H.J. Kushner and P.A. Whiting. "Asymptotic Properties of Proportinal-Fair Sharing Algorithms: Extensions of the Algorithm". In *Proc. 2003 Allerton Conf.*, Urbana, IL. Univ. of Illinois Press.

[6] H.J. Kushner and G.G. Yin. *Stochastic Approximation Algorithm and Applications.* Springer-Verlag, 1997.

[7] J. Lee, R.R. Mazumdar, and N.B. Shroff. "Opportunistic Power Scheduling for Multi-server Wireless Systems with Miimum Performance Constraints". *IEEE INFOCOM*, 2004.

[8] A.L. Stolyer. "On the Asymptotic Optimality of the Gradient Scheduling Algorithm for Multi-User Throughput". 2003.

[9] D. Tse. "Downlink Scheduling for Qualcomm's HDR". 1999.